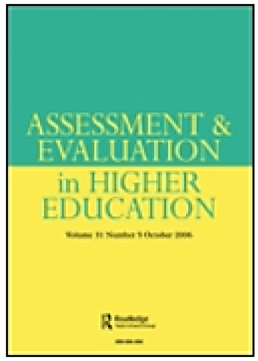

Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics

Journal:	*Assessment and Evaluation in Higher Education*
Manuscript ID	Draft
Manuscript Type:	Original Paper
Keywords:	learning analytics, agency, equity, transparency

Taylor and Francis Word Template for journal articles

Yi-Shan Tsai[a] [0000-0001-8967-5327]*, Carlo Perrotta[b] [0000-0003-3572-0844] and Dragan Gašević[c] [0000-0001-9265-1908]

[a]*School of Informatics, University of Edinburgh, Edinburgh, United Kingdom;* [b]*Faculty of Education, Monash University, Clayton, Australia;* [c]*Faculty of Education, Monash University, Clayton, Australia*

yi-shan.tsai@ed.ac.uk

Provide short biographical notes on all contributors here if the journal requires them.

Empowering learners with personalised learning approaches? Agency, equity and transparency in the context of learning analytics

Abstract

The emergence of personalised data technologies, such as learning analytics (LA) is framed as a solution to manage the needs of higher education student populations that are growing ever more diverse and larger in size. However, the current approach to learning analytics presents tensions between increasing student agency in making learning-related decisions and 'datafying' students in the process of collecting, analysing and interpreting data. This paper presents a study that explores staff and student experience of agency, equity, and transparency in existing data practices and expectations towards LA in a UK university. The results show a number of intertwined factors that have contributed to the tensions between enhancing a learner's control of their studies and, at the same time, diminishing their autonomy as an active agent in the process of LA. This paper argues that learner empowerment should not be automatically assumed to have taken place as part of the adoption of learning analytics. Instead, the interwoven power relationships in a complex educational system and the interactions between humans and machines need to be taken into consideration when presenting LA as an equitable process to enhance student agency and educational equity.

Keywords: learning analytics, agency, equity, transparency

Introduction

In education, the trend towards data-based methods of governance and management initially led to a thriving field of educational data mining, concerned with the automated exploration of data from educational settings (Siemens and Baker 2012). Later, it enabled the emergence of Learning Analytics (henceforth LA) as a distinct field of research and practice that aims to use data to optimise learning and the environments where it occurs (Long et al. 2011). While LA shows great potential in tackling educational challenges of attainment and student retention, there are also prevailing concerns around the ethical and privacy implications of the use of student data, and the extent to which LA can benefit every student (Tsai, Moreno-Marcos, Jivet, et al. 2018).

A key problem with current approaches to LA lies in the tensions between increasing student agency in making learning-related decisions and 'datafying' students in the process of collecting, analysing and interpreting data. On the one hand, digital technologies in education have been associated with the rise of a distinct form of market-based individualism, which shifts the traditional values of education from public good to private interest (Castañeda and Selwyn 2018). This frames learning as a self-centred endeavour and investment (Thompson and Cook 2017), with learners entrusted with more responsibility to improve their own performance through technology-enhanced support. On the other hand, the indiscriminate collection and analysis of student data from digital learning environments risks disregarding human factors and the socio-cultural contexts in which the data is generated (Perrotta 2013; Gašević, Dawson, and Siemens 2015). These tensions bring to the fore an urgent need for a critical discourse to further examine the paradoxical promises of LA in enhancing student agency, while furthering a pervasive governance culture of data collection, interpretation, and intervention design, thereby contentiously exercising 'algorithmic control' over education.

In this paper, we argue that learner empowerment should not be automatically assumed to have occurred through the adoption of personalised data technologies such as LA. Instead, the interwoven power relationships in a complex educational system and the interactions between humans and machine need to be taken into consideration when presenting LA as an equitable process to enhance student agency and educational equity. We reflect on the aforementioned issues by drawing on data collected from six student focus groups (26 participants in total) and 5 staff focus groups (18 participants in total) carried out in a UK higher education institution. The analysis was informed by two research questions:

1. How might personalised data technologies enhance or hamper equity and agency?
2. How might existing and expected transparency of data practices strengthen or compromise student agency?

With the first question, we intend to draw attention to the prevailing assumption that personalised data technologies can empower learners (Kurilovas, Krikun, and Melesko 2016; Mouri et al. 2016; Charleer et al. 2018). This assumption is widespread, despite some evidence suggesting that presenting students with their own data can have a negative impact on motivations and chances of academic success (Lonn, Aguilar, and Teasley 2015). In line with this more critical research, we examine this assumption from the aspects of equity and agency in the context of LA. The second question explores the presence of student agency by looking into the control of their own data. In particular, we present issues on information asymmetries between data collectors and data subjects due to power imbalance (Acquisti and Grossklags 2007; Rubel and Jones 2016). This paper critically examines the extent to which learning analytics can be used to enhance student agency and educational equity.

In the next section, we discuss key concepts of datafication, agency, equity and transparency, drawing on relevant contributions from various disciplinary perspectives, including critical sociology. Thereafter, we present the participants' experience of agency, equity, and transparency in existing data practices and their expectations of LA. In the conclusion, we outline the problems to address when implementing LA as a means to create a more inclusive and equitable learning environment in higher education.

Agency and transparency in the context of learning analytics

Education in a data-led society

Much has been written about the role of data and algorithms in society, and several perspectives from diverse theoretical orientations are now available (Hallinan and Striphas 2016; Kelling et al. 2009; Kitchin 2014; Turow, McGuigan, and Maris 2015). The ongoing multidisciplinary debate is concerned with the extent to which Big Data enables novel ways of understanding the world and acting upon it. Drawing on the literature, Kitchin (2014) argues that Big Data are often defined as huge in volume, high in velocity, diverse in variety, exhaustive in scope, striving to capture entire populations (N=all), relational in nature, flexible and scalable. Big Data is therefore qualified using attributes evoking power and comprehensiveness, but also heterogeneity and uncertainty. As Kitchin notes (2014, 2):

> The challenge of analysing Big Data is coping with abundance, exhaustivity and variety, timeliness and dynamism, messiness and uncertainty, high relationality, and the fact that much of what is generated has no specific question in mind or is a by-product of another activity.

This has led to an assertive, optimistic form of empiricism underpinned by a presumed ability of analytical approaches to generate new insights from Big Data that partial, sampled datasets cannot guarantee. Indeed, this was the spur behind several forms of data analytics in specific domains, with LA being no exception.

In particular, LA has emerged as a solution to address prevalent challenges in education, such as student retention, widening access, and personalised support for a massive student cohort (Ferguson 2012; EDUCAUSE 2018). The two main aims of LA are the diagnosis and prediction of various dimensions of educational performance, both geared towards the production of 'actionable insights' of immediate and demonstrable instructional effectiveness (Clow 2013; Siemens 2013). Other popular trends include using LA to provide personalised feedback at scale

(Pardo et al. 2019) and to identify variables and behaviours that promote student success and address the need for quality assurance of educational services (Lester et al. 2018). Theoretically, the field of LA is broadly aligned with scholarship in the learning sciences, assessment and instructional design – while simultaneously positioning itself as a collection of computational innovations (mostly from data science), made possible by the growing penetration and ubiquity of digital platforms and devices in education. Similar to themes in the Big Data discourse, LA is susceptible to the enduringly partial nature of whole datasets, which remain shaped by the contingent sociotechnical conditions in which they are generated, the dependence on using technologies for measurement, storage and digitisation, as well as the contextual and domain-specific assumptions that underpin the deployment of computational methods (Kitchin 2014). These constraints impose questions on the degree to which learning analytics can present faithful and fair information about learners in different disciplines and from different socio-cultural and economic backgrounds.

In the context of education, equity has two dimensions. The first is fairness, which ensures opportunities to achieve personal potential without being impeded by personal conditions. The second is inclusion, which ensures a basic minimum standard of education for all (Simon, Kuczera, and Pont 2007). Similar themes have been highlighted in some LA circles, with several authors expressing concerns about the unfair differential impact of predictive models in education (Prinsloo and Slade 2014; Roberts, Chang, and Gibson 2017). This interest in the fairness of LA is, of course, a reflection of a broader social and scientific debate about the dangerous tendency of predictive modelling to reproduce existing biases based on race, gender and class (Richardson, Schultz, and Crawford 2019). Indeed, predictive fairness is an emerging area of experimentation in the LA field, with new promising techniques such as 'slicing analysis' (Gardner, Brooks, and Baker 2019) being proposed. Slicing analysis evaluates model accuracy for different sub-groups or individuals to identify unfair differences, which can be used to identify fairer model. However, questions remain about the tendency to treat fairness and justice as properties of computational models, rather than properties of social systems (Selbst et al. 2019). This means that, for example, innovations to make analytics-based predictions in a MOOC 'fairer' might miss the point if they fail to acknowledge the broader conditions that make MOOC participation more likely among specific gender or race groups.

Agency and data in education

The pervasiveness of digital technology has inspired a public debate about fundamental aspects of human nature (Castañeda and Selwyn 2018). Critical arguments within and beyond academia often take aim at data-based surveillance, algorithmic manipulation of behaviours and artificial intelligence to ask rather philosophical questions about what it means to be human, or to have 'agency'. While an in-depth examination of this topic is beyond the scope of our article, it is important to identify some key ideas that have particular relevance to the present study, i.e. how human agency is constrained by digitisation and automation.

In research that focuses on formal educational settings, i.e. schools and higher education, agency is generally understood as cognitive and metacognitive agency: a collection of active processes of knowledge acquisition or development, as well as the complex assortment of individual strategies that allow awareness of and control over those processes. A number of studies in the learning sciences and psychology have explored various aspects of agency in relation to human cognition (computational aspects, situatedness, schemata, motivations, dispositions and so forth) from a relatively individualistic point of view (Nicholls 1984; Dweck 2000). In education, this is translated into an emphasis on students as rational agents with the potential to take responsibility for their own learning (Crick and Goldspink 2014). This has had a notable influence on the development of LA as a distinct discipline (Shum and Crick 2012). Indeed, aspects of individual cognition have been computationally modelled and then used for the development of various forms of adaptive or AI-enabled instruction, including various flavours of 'personalised learning', such as intelligent tutoring systems (providing instructional advice on a one-to-one basis, akin to human tutors), recommender systems (predicting a user's preference or needs for an item), and pedagogical agents (simulated figures designed to facilitate interactions between learners and the computer programme). The individualistic slant of a large part of research in the learning sciences (and by extension in LA) reflects traditional empirical foci in psychology.

By contrast, sociological perspectives tend to favour a different view where human agency is profoundly shaped by structural factors. In the broader context of digital innovation and the so-called 'datafication' trend, these sociological perspectives have produced a number of critical arguments, several of which attend to educational technology and LA in particular. An increasingly vocal debate within this 'camp' discusses agency in algorithmic 'systems of control' (Agre 1994; Kitchin and Dodge 2012; Williamson 2015); that is, systems where computational power is a tool in the service of a pervasive culture of governance that seeks to exert control through economic rationality, efficiency and individual accountability . This culture is seen as the result of global factors and influences, which contributed to derail the process of digital innovation away from the emancipation and the empowerment of human agency, and towards compliance, control and, often, outright surveillance. In the context of LA, these arguments have translated into a critique of key notions such as 'actionable intelligence' as the focus of LA has shifted, according to some (e.g. Knox 2017), from hindsight to foresight and prediction. The emphasis on action informed by predictive models has, for its critics, a tendency to prioritise effects and indicators (signals) over causes, thus leading to narrow remedial strategies in which students and teachers are channelled along predefined trajectories of educational performance that, paradoxically, leave little room for agency.

Transparency and data in education

Transparency is a major theme in current discussions about algorithmic accountability (e.g. Ananny and Crawford 2018; Tsai and Gašević 2017). This debate has important implications for LA. Indeed, transparency is one of the overarching goals of the LA project, which seeks to make learning visible and measurable in order to inform actionable feedback. The transparency theme is somewhat reversed in critical arguments while retaining the theme of empowerment through openness. Here, the emphasis is on the need to make computational systems more accountable in relation to the collection and manipulation of personal data: black boxes to be opened up and critically interrogated (Pasquale 2015). In both cases, the underlying assumption is that positive outcomes (evidence-based learning and democratic accountability) will be attained by rendering complex realities more transparent.

Transparency and accountability are treated as preconditions for the production of authenticity: authentic learning, or authentic democratic accountability. This approach may lead to a number of issues, including what Ananny and Crawford (2018, 7) call a 'false binary' between complete secrecy and total openness. On the one side, complete secrecy is not only unattainable, but problematic in its own right in several institutional settings, including education; on the other, total openness mistakenly assumes that individuals are informed, rational agents perfectly positioned to maximise benefits from publicly available information. As such, the rhetoric of transparency in all its manifestations – i.e. as the desirable outcome of analytics or as an ethical imperative for algorithmic methods – may privilege 'seeing over understanding' (*ibid*., 8). It could be argued that the prioritisation of transparency as visibility over self-reflexive knowledge also underpins the current political discourse of institutional disclosure. For example, the new European General Data Protection Regulation (The European Parliament and the Council of the European Union 2016) intends to empower individuals with the right and responsibility to make decisions regarding the use of their personal data, while institutions are held accountable for ensuring the transparent provision of relevant information to enable this process. However, the inherent imbalance in the power relationships in the various contexts in which data is collected poses questions about the extent to which individuals can truly make informed decisions about the use of their data. The implementation of student-centred learning analytics is no exception (Knox 2017).

Methodology

This study sets out to understand teachers' and students' expectations of LA, and questions the extent to which LA can empower learners and enhance equity in education, so as to move towards a deliberative, democratic integration of LA. A focus group was chosen to capture data, taking advantage of dynamics in a group where participants inspire one another and probe ideas among themselves (Liamputtong 2011) to increase data richness. In particular, focus groups allow for shared experiences among the participants that increases willingness to discuss personal views.

Participants

To enable in-depth discussions, a focus group typically involves a relatively small number of participants ranging from four to twelve (Liamputtong 2011). This study involved six student focus groups, each comprising four to five participants. Participants were invited widely from a comprehensive university in the UK to include a diversity of student bodies from different disciplines and degree types. We received 139 positive responses from students, and we selected six students for each group (6 groups) to represent as many different disciplines as possible. The selection process was first-come-first-serve, with the constraint that, where possible, participants were chosen from differing disciplines. However, only 26 students (7 males, 19 females) participated in this case study at the end due to late withdrawals or absence. The six groups are labelled as UG1, UG2, UG3, UG4, PG, and ODL in this paper. UG1 and UG2 comprise undergraduate students from the Arts, Humanities, and Social Sciences College, which had the largest student body compared to the other colleges. UG3 comprise undergraduate students from the Science and Engineering College, and UG4 from the Medicine and Veterinary Medicine College. PG includes postgraduate students from mixed disciplines and ODL consists of online-distance learning students from mixed disciplines. Only one participant from the ODL group had past experience with learning analytics.

The process of sampling for teaching staff focus groups proved challenging partly due to their busy work schedules. Five focus groups of participants were sampled widely from the three university colleges mentioned above (labelled as G1 to G5 in this paper). Twenty-five teaching staff volunteered but only eighteen (10 males, 8 females) managed to attend the focus groups.

As a result, three of the groups comprised three to five participants respectively, and two comprised only two participants due to late withdrawals or absence. Five of the participants had director roles (e.g., programme director or director of undergraduate studies), and three had personal tutor roles. Not all the participants had experience with LA, although most of them had a certain degree of experience of working with or using data to inform their teaching practices.

Procedure

The focus group interviews were semi-structured, each lasting approximately an hour. All participants received a short introduction to the concept of learning analytics before the focus group interviews started. As the institution's adoption of learning analytics was at a rather early stage, the focus groups were intended to understand participants' awareness and attitudes regarding existing data practices, which the interviewer drew upon to guide participants to consider the potential benefits and challenges of using student data for learning analytics. To this end, ten different questions were designed for staff and student focus groups respectively to understand their current experience with existing data practices at the university and expectations or desires to address learning and teaching challenges through LA (accessible here: http://bit.ly/FG_questions). All participants signed a consent form to participate in the study and agreed to have their conversations recorded. Each student received ten pounds and each teaching staff member received lunch in gratitude for their time.

A thematic analysis

The focus group interviews were transcribed verbatim and then analysed using a thematic coding method (Grbich 2012). The coding scheme was developed inductively, which involved the researcher reading the transcripts repeatedly to identify recurring themes and types of issues raised. The qualitative analysis tool – Nvivo – was employed to assist in this process. In total, 64 codes categorised into 3 main themes and 14 sub themes were developed to analyse student focus groups (accessible here: http://bit.ly/students_coding), while 59 codes categorised into 4 main themes and 26 sub themes were developed to analyse staff focus groups (accessible here: http://bit.ly/staff_coding).

In the following sections, the student participants are denoted as S (student) and teacher participants as T (teachers) with numbers (1 to 5) to differentiate between individuals in the same group. Some of the participants were second language speakers of English. The selected excerpts are faithful to the original responses, with the minor exception that some redundant words, such as 'like' and 'you know', were edited out whenever these words were not considered to contribute significant meaning.

Results and discussion

Our engagement with teaching staff and students highlighted the existence of different perspectives regarding the role of LA in enhancing equity, agency and transparency, with some notable misalignments in expectations. These differences need to be considered carefully when higher education institutions deploy LA so as to cultivate a sense of ownership. In this section, we present the results in accordance with the two research questions introduced previously.

How might personalised data technologies enhance or hamper equity and agency?

Personalisation and equity

Learning analytics promises to tailor support to individuals by profiling students using their learning data and demographic characteristics so as to devise a suitable intervention. In this way, learning analytics strives to help individuals achieve their optimal potential rather than bringing every student to the same level of performance. This personalised approach demonstrates potential in enhancing equity by acknowledging that education is by no means one-size-fits-all and students at different learning stages require different levels of support. However, targeted support arguably risks labelling certain groups of students while seemingly disadvantaging other students by directing resources away from them. We highlighted in our literature review that fairness and equity are emerging as important concerns in LA, but caution is needed when treating these dimensions as properties of models rather than properties of social systems. Our qualitative data extends this point further, providing an insight into the contextual and personal factors that a 'diverse' approach to LA could engage with. For example, some student participants suggested that LA should help instructors and programme directors understand the educational backgrounds and learning needs of different students, so as to provide relevant support instead of overloading students with superfluous information:

> I have ASD [autism spectrum disorder] which causes me social and communication issues and I've sort of spent the first year and a half being quite lonely and isolated at University and yet overwhelmed by lots of information, but unable to sift through it

> [information bundled up in a generic information] and work out exactly what was of benefit specifically to me (S4, UG1).

> If English is not your first language, they could use that and make you aware of the support that is available to you as a foreign language speaker. Now there were some people who were from very unprivileged backgrounds or people who lived at home and didn't live at a university accommodation. For them you could be giving more guidance of how to get involved. So just looking at your background and sending an email to the people saying, 'Look, this is available'. (S3, UG1)

While the students generally believed that a personalised approach to educational offerings and learning support can enhance their educational achievement, some students pointed out the problem of unfairness when restricting access to resources under the banner of personalisation:

> For the personalisation of services, I wouldn't go so far, because we want to have equal access to all the services at the university. I would give them data about my high school years or what curriculum I studied to assess whether all students are on the same starting level, maybe the beginning of university. And for those purposes to sort of tailor initial support services (S3, UG2).

Similarly, the teaching staff expressed a desire to understand the nature of the student body and relationships between learning behaviour and performance, so as improve the planning and delivery of a course according to the needs of a growing population of students from diverse cultural and educational backgrounds:

> I think that sometimes these processes of using data and things like that can help us with this [widening access]. You know, they can help us to understand the nature of our student body more effectively and try to tune the ways that we work with them more effectively (T3, G3).

However, they were also concerned about 'fairness' in targeted support and the pedagogical effectiveness:

> I suppose there could be an argument for equity of treatment. You're taking a particular class of students and you're putting much more effort into them than the rest. We can sort of say that part of education is being given the freedom to fail on your own, as opposed to school, you sort of learn from that (T4, G2).

Despite the common interest among students and teaching staff in using LA to improve curriculum design and student support, teaching staff tended to look for the big picture of a student cohort whereas students focused on the differences between individuals and expected educational equity to be achieved by optimising everyone's opportunity to excel. Nevertheless, fairness emerges as a central concern around personalisation for both students and teachers. Moreover, it poses a paradoxical question concerning whether personalisation enhances one's opportunity to succeed or takes away the opportunity to learn from failures.

Personalisation and agency

Agency is the capacity of individuals to 'control' and 'compose' their behaviour for a determined end, and to anticipate how others would interpret their behaviour (Enfield and Kockelman 2017). In the context of education, agency is characterised by choice-making in learning; generating new knowledge; taking responsibility for learning; and engaging in learning relationships (Crick and Goldspink 2014). LA aims to promote student agency by positioning students as active actors to make data-informed decisions related to learning (Kurilovas, Krikun, and Melesko 2016; Mouri et al. 2016; Charleer et al. 2018), such as adjusting learning strategies upon critical reflections of their behavioural patterns or performance. However, the range and amount of data that can be collected for learning analytics to identify suitable support for individuals has often led to ethical issues around surveillance and 'datafying' students (Zuboff 2015). In the focus groups, students, while agreeing that learning was their own responsibility, expected personalised feedback to guide them in making learning- or career-related choices. In contrast, teaching staff highlighted a concern about suppressing learner autonomy through excessive support.

For example, the students pointed out that information about their learning progress could help them spend time more efficiently by focusing on areas that need to be improved or working strategically towards the next assignment or exam. This was perceived as especially beneficial in the early years of higher education when students are still trying to adapt to a learning mode that involves fewer interactions with instructors, but more independent effort on the learner's side.

> Ultimately it is, it's our responsibility. We're adults. We're in control of our own learning but that doesn't mean that support and guidance and help and pastoral care aren't still important especially in the first few years (S4, UG1).

Another student pointed out the struggle when being encouraged to explore 'whatever they want to do':

> Sometimes freedom just makes you lost…. I think what is more helpful is they [instructors] really show us the optimum way to get to our career rather than let us do whatever we want, 'cause they know how we've been performing all these years, all these semesters, and they can see what is our opportunities (sic) in certain areas (S1, CAHSS2).

To the students, personalised support and guidance can scaffold the process of exerting agency on one's own learning decisions. They objected to the idea of framing education as a self-centred endeavour, as technology-based learning has increasingly been positioned as (hyper)individualised or less collective (Castañeda and Selwyn 2018). Similarly, the teaching staff being interviewed agreed that LA has the potential to enhance student agency with appropriate support to help them understand and interpret data, thereby leading to self-initiated changes in behaviours. However, several participants also raised concerns that learning analytics could potentially hamper rather than enhance student agency:

> The more we start identifying individual students, 'well, you need a remedial class because you're underperforming', you're kind of taking that agency away from students. And I think there is a very big danger of this kind of approach…. Spoon feeding students, telling them what they have to know, giving them sort of tests and stuff, has been the way that universities responded to poor satisfaction scores, poor teaching scores, or whatever it is. In other words, instead of saying 'students, listen, we need a dialogue about this', it's been more prescriptive action (T4, G3).

> The one thing that we must get over to our students is the primary responsibility for their learning is there. That really is the bottom line. I mean that's what you do when you leave [the university]…. What you're telling the employer [is] this person can cut the mustard (T2, G2).

A comparison of the responses from the students and staff reveals mismatched interests among the key stakeholders of LA. Students tended to focus on addressing their current struggles, while institutions focused on responding, perhaps in a rather haphazard manner, to the results of student satisfaction surveys. By contrast, the teaching staff were more concerned that unbalanced (constant, excessive, or dictating) personalised interventions can have negative impacts on the development of a student's problem-solving abilities, which is sometimes gained through a painful learning process. It was notable that the students emphasised the need for personalised support to enhance *choice-making in learning* (Crick and Goldspink 2014), whereas the teaching staff were more concerned with helping students to develop a sense of *responsibility for learning* (*ibid.*).

Some teaching staff pointed out the paradox of LA in promoting student agency with targeted support and meanwhile diminishing it through constant surveillance in online learning environments:

> There is a series of parallel demands that actually play against one another: being more independent, having more freedom, and they are being monitored much more closely… (T2, G4).

Moreover, the issue of surveillance in creating a sense of remoteness and distance between data subjects and the data collector (Bauman and Lyon 2012) has also led to discussion on 'being treated as numbers' among the student focus groups:

> See that this person is beyond just data…. not reducing a person just to the figures that are being shown on your laptop regarding the person's performance…. You have to understand why the numbers are coming…. I feel like interaction is the key…to understand the data you need to understand where it's coming from (S4, UG2).

Aligned with the students' views, several teaching staff highlighted the risk of removing human factors by discounting the professional knowledge of teachers or decontextualizing data that are produced by students who each have different personal circumstances and learning approaches:

> I don't want it [LA] to make all of the students behave in the exact same way to satisfy an algorithm. I want it to enable students to have the best experience in whatever that experience is. You know, you can be totally different from everyone else and still do perfectly fine. I want it [LA] to…enable students to do better and not make them all mini 'me's (T2, G5).

Here, we observed resistance to the algorithmic control that has been pervasively used to enhance economic efficiency in educational contexts (Williamson, 2015), and a call to reflect on how technologies mould people's emotional and cognitive interactions with each other and with the machine (Castañeda and Selwyn 2018).

How might existing and expected transparency of data practices strengthen or compromise student agency?

Although the students being interviewed were generally aware that the institution collected and used certain types of data about them, such as academic and immigration and study permit data, it was not clear to them who could access the data and how it could be used to improve academic offerings and support. In general, a phenomenon of information asymmetry was observed, which arguably diminishes student agency in giving informed consent about the use of their data. Firstly, the implications of consent given to have one's data collected were not clear to students at the time of enrolment:

> You have to agree to share this data otherwise you wouldn't enrol, so you are not probably thinking that much about the consequences of every single piece of data that you provide to the university. It's just because it's a part of the [application] process (S4, UG3).

> I've only just sort of begun my studies so I don't think I have enough time in to say 'well, I don't think you should have collected that' (S1, ODL).

> I would assume that [I gave consent at enrolment]. I think often you're signing consent for things you don't realise…. (S2, PG)

Although the phenomenon of exchanging data for education is likely to have changed since the implementation of 2016/679 GDPR (The European Parliament and the Council of the European Union 2016) in May 2018, the 'lawfulness of processing' in GDPR still allows institutions to process personal data when it is necessary for the purpose of 'legitimate interests' or to carry out tasks that are of 'public interest'. Moreover, although these students were explicitly asked to provide consent, the priority to complete the enrolment process at the moment they came to the university was likely to cloud their risk assessment on data sharing, or to lead to indifference in the consequences of sharing personal data. As a result, students compromised on consent-giving out of *rational ignorance* – users consider the effort and loss of time in reading a lengthy and complex policy to outweigh the perceived risk of disclosing personal information (Acquisti and Grossklags 2007). Our analysis provides some evidence in support of contributions that critique notions of 'ideal' transparency, which places the burden to seek out information about a system on individuals, and reinforces the mistaken assumption that people will hold perfect information to make rational decisions and give fully informed consent (Ananny and Crawford 2018).

Secondly, there has been insufficient communication between the institution and students regarding what happens after data is collected. The involvement of students often ceases after their data has been collected. This results in limited understanding of the benefits of opting into data collection, especially when it comes to giving feedback to improve teaching and educational services in general:

> Sometimes it just feels like they kind of have to take feedback at the end of the courses just because it looks nice, but you don't really know if they actually even read it or use it (S3, UG3).

> It is always like individual feedbacks between us and the university and then we don't know what the university do, and the university decides by themselves. We are being separated (S1, UG2).

This indicates that when opt-in or opt-out options are made available to students sharing data in exchange for particular LA services or interventions, concrete examples showing how the loop from data collection to action is closed are crucial to informed consent. This also applies to cases where service providers are involved, as the students in general showed distrust in sharing data externally for the fear of becoming the targets of commercial advertisements.

Thirdly, cognitive limitations in understanding the algorithms embedded in LA systems could compromise the rational choices in sharing personal data, known as the phenomenon of *bounded rationality* – self-disclosure decisions are not always rational due to perceived or actual cognitive limitations in understanding all necessary information required to give informed consent (Acquisti and Grossklags 2007). In some cases, the opaqueness of algorithms results in distrust of analytics. This point in particular was raised by teaching staff:

> The cleverer the algorithm, the opaquer and therefore the more dangerous it is…. We don't know what biases are actually built into the data because the way in which the data are gathered are contaminated by, for example, issues of race and gender and so on. So it doesn't take long before an algorithm becomes magic. It becomes something beyond our understanding and that's dangerous (T1, G5).

> We're kind of caught between two competing demands from the same group [students], which is on the one hand the demand for more transparency and more information, and on the other hand, the fact that oftentimes the outcome of that is not

> in any explicable or direct way correlated to the results or what they're doing in the course. And it actually produces an effect that we don't realise and they don't realise as well (T1, G4).

In this case, information asymmetry results from a power imbalance caused by humans making decisions based on second-hand information (selective summary statistics provided by algorithms). The complexity of a technological system poses another limitation of transparency; that is, seeing does not necessarily lead to understanding (Ananny and Crawford 2018). Moreover, the opaqueness of algorithms can dangerously direct attention away from the process of learning activities and the associated social, cultural, and political factors in the broader context. As a result, it becomes an almost impossible task for students to challenge the precision of analytics of their learning.

Conclusion

Our analyses have identified conflicting interests in LA among students, teachers, and institutions, which have led to mismatched perceptions and expectations about the way LA can be used to enhance learner agency and improve equity. We argue that LA-based interventions should not be assumed to empower students. A number of intertwined factors have contributed to the tensions between enhancing a learner's control of their studies and, at the same time, diminishing their autonomy as an active agent in the process of LA. These factors include the way interventions are devised and delivered, the way data is collected, analysed, and interpreted, the transparency of the data process, and the opaqueness of algorithms. We summarise our findings with three recommendations to mitigate the observed tensions:

First, interventions need to be based on the learning sciences to balance what students want and what is good for them. It is notable that students are concerned about making good decisions regarding learning and career development, whereas teachers highlighted the risks of learning analytics in terms of spoon-feeding students, leading to the removal of agency. While these two views do not necessarily conflict with each other, they highlight the importance of devising LA-triggered interventions based on learning sciences. For example, the seven principles of feedback proposed by Nicol and Macfarlane-Dick (2006) could be used as a framework for LA-based feedback to drive self-regulation.

Secondly, LA needs to leverage rather than replace human contact. Key to this is a realistic evaluation of staff capacity and capability to deliver interventions. While the students generally bought into the rhetoric of personalisation advanced by learning analytics, their idea of personalisation was not necessarily a computational one. In fact, they lamented the lack of individualised support from human tutors, while showing a degree of suspicion for automated systems that focus on coarse, dichotomous metrics of educational performance, such as the risk of dropping out or failure. On the one hand, this suggests a desire to maintain a degree of control, indeed agency, over learning; on the other, it points to a notion of support that is dialogic, human-based and, inevitably, labour intensive in nature. Mirroring these concerns, teachers pointed to the confusion between individualised support and 'industrial-scale provision' – a confusion that has been introduced by the institution as a result of pressures to widen access to higher education as well as demonstrate performance via quantitative indicators, such as student satisfaction and progression. As a result, the version of personalisation expected by students was considered unrealistic by teaching staff, in that such levels of support cannot be delivered by humans without placing undue pressures on already heavy workloads.

Thirdly, issues of transparency and visibility in terms of data policies, practices, and algorithms, requires a more informed debate around the implications for agency. An important question, in this regard, is the following: to what degree does agency depend on a relative lack of visibility and transparency? Despite the fact that obtaining explicit consent before data is collected and processed has been acknowledged as a requirement in Europe, the ineffective communication of policies puts burden on students seeking to comprehend the consequences of giving data away and being responsible for it. This appears like an effort to fulfil an obligation rather than to help students develop agency in any constructive way. Moreover, the disengagement of students in phases beyond data collection and the challenge of making algorithms transparent in a comprehensible way discredit the idea that LA empowers students, as little room has been left for agency in this remedial approach.

The paradox of agency highlights the need to 'deobfuscate' the politics of data-based personalisation. Indeed, this is not only an ethical priority but also a methodological one, concerned with a more accurate understanding and communication of the inner 'sociality' of algorithmic diagnosis and prediction. It is crucially important to acknowledge and address the conflicting beliefs about data-based personalisation, surveillance, and agency when introducing LA as an equitable solution to educational challenges. In order to mitigate these conflicts, institutions need to intentionally involve different groups of users in a partnership to design and implement LA (Dollinger and Lodge 2018; Roberts, Chang, and Gibson 2017), and develop a context-based policy (Tsai, Moreno-Marcos, Tammets, et al. 2018) that ensures the deployment of LA to align with the institutional values of inclusion, equity, and student autonomy.

References

Acquisti, Alessandro, and Jens Grossklags. 2007. "What Can Behavioral Economics Teach Us about Privacy?" In *Digital Privacy: Theory, Technologies, and Practices*, edited by Alessandro Acquisti, Stefanos Gritzalis, Costos Lambrinoudakis, and Sabrina di Vimercati, 363–77. CRC Press.

Agre, Philip E. 1994. "Surveillance and Capture: Two Models of Privacy." *The Information Society* 10 (2): 101–27. https://doi.org/10.1080/01972243.1994.9960162.

Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20 (3): 973–89. https://doi.org/10.1177/1461444816676645.

Bauman, Zygmunt, and David Lyon. 2012. *Liquid Surveillance: A Conversation*. 1 edition. Cambridge, UK ; Malden, MA: Polity Press.

Castañeda, Linda, and Neil Selwyn. 2018. "More than Tools? Making Sense of the Ongoing Digitizations of Higher Education." *International Journal of Educational Technology in Higher Education* 15 (1): 22. https://doi.org/10.1186/s41239-018-0109-y.

Charleer, Sven, Andrew Vande Moere, Joris Klerkx, Katrien Verbert, and Tinne De Laet. 2018. "Learning Analytics Dashboards to Support Adviser-Student Dialogue." *Ieee Transactions on Learning Technologies* 11 (3): 389–99. https://doi.org/10.1109/TLT.2017.2720670.

Clow, Doug. 2013. "An Overview of Learning Analytics." *Teaching in Higher Education* 18 (August): 683–95.

Crick, Ruth Deakin, and Chris Goldspink. 2014. "Learner Dispositions, Self-Theories and Student Engagement." *British Journal of Educational Studies* 62 (1): 19–35. https://doi.org/10.1080/00071005.2014.904038.

Dollinger, Mollie, and Jason M. Lodge. 2018. "Co-Creation Strategies for Learning Analytics." In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 97–101. LAK '18. New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170372.

Dweck, Carol S. 2000. *Self-Theories: Their Role in Motivation, Personality, and Development*. Psychology Press.

EDUCAUSE. 2018. "NMC Horizon Report Preview: 2018 Higher Education Edition." NMC Horizon Project. https://library.educause.edu/resources/2018/4/nmc-horizon-report-preview-2018.

Enfield, N. J., and P. Kockelman, eds. 2017. *Distributed Agency*. New York, NY: Oxford University Press.

Ferguson, Rebecca. 2012. "Learning Analytics: Drivers, Developments and Challenges." *International Journal of Technology Enhanced Learning* 4 (5/6): 304. https://doi.org/10.1504/IJTEL.2012.051816.

Gardner, Josh, Christopher Brooks, and Ryan Baker. 2019. "Evaluating the Fairness of Predictive Student Models Through Slicing Analysis." In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 225–234. LAK19. New York, NY, USA: ACM. https://doi.org/10.1145/3303772.3303791.

Gašević, Dragan, Shane Dawson, and George Siemens. 2015. "Let's Not Forget: Learning Analytics Are about Learning." *TechTrends* 59 (1): 64–71. https://doi.org/10.1007/s11528-014-0822-x.

Grbich, Carol. 2012. *Qualitative Data Analysis: An Introduction*. SAGE.

Hallinan, Blake, and Ted Striphas. 2016. "Recommended for You: The Netflix Prize and the Production of Algorithmic Culture." *New Media & Society* 18 (1): 117–37. https://doi.org/10.1177/1461444814538646.

Kelling, Steve, Wesley M. Hochachka, Daniel Fink, Mirek Riedewald, Rich Caruana, Grant Ballard, and Giles Hooker. 2009. "Data-Intensive Science: A New Paradigm for Biodiversity Studies." *BioScience* 59 (7): 613–20. https://doi.org/10.1525/bio.2009.59.7.12.

Kitchin, Rob. 2014. “Big Data, New Epistemologies and Paradigm Shifts.” *Big Data & Society* 1 (1): 2053951714528481. https://doi.org/10.1177/2053951714528481.

Kitchin, Rob, and Martin Dodge. 2012. *Code/Space: Software and Everyday Life*. MIT Press. https://mitpress.mit.edu/books/codespace.

Knox, Jeremy. 2017. “Data Power in Education: Exploring Critical Awareness with the ‘ Learning Analytics Report Card.’” *Television & New Media* 18 (8): 734–52. https://doi.org/10.1177/1527476417690029.

Kurilovas, E., I. Krikun, and J. Melesko. 2016. “On Using Learning Analytics to Personalise Learning.” *ICERI2016 Proceedings*, 6987–96.

Lester, Jaime, Carrie Klein, Aditya Johri, and Huzefa Rangwala. 2018. *Learning Analytics in Higher Education: Current Innovations, Future Potential, and Practical Applications*. Routledge.

Liamputtong, Pranee. 2011. *Focus Group Methodology: Principle and Practice*. Sage Publications.

Long, P. D., G. Siemens, G. Conole, and D. Gašević, eds. 2011. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK’11)*. Banff, AB, Canada: ACM. https://tekri.athabascau.ca/analytics/.

Lonn, Steven, Stephen J. Aguilar, and Stephanie D. Teasley. 2015. “Investigating Student Motivation in the Context of a Learning Analytics Intervention during a Summer Bridge Program.” *Computers in Human Behavior* 47: 90–97. https://doi.org/10.1016/j.chb.2014.07.013.

Mouri, Kousuke, Hiroaki Ogata, Noriko Uosaki, and Erdenesaikhan Lkhagvasuren. 2016. “ Context-Aware and Personalization Method Based on Ubiquitous Learning Analytics.” *Journal of Universal Computer Science* 22 (10): 1380–97.

Nicholls, John G. 1984. “Achievement Motivation: Conceptions of Ability, Subjective Experience, Task Choice, and Performance.” *Psychological Review* 91 (3): 328–46. https://doi.org/10.1037/0033-295X.91.3.328.

Nicol, David J., and Debra Macfarlane-Dick. 2006. “Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice.” *Studies in Higher Education* 31 (2): 199–218. https://doi.org/10.1080/03075070600572090.

Pardo, Abelardo, Jelena Jovanovic, Shane Dawson, Dragan Gašević, and Negin Mirriahi. 2019. “Using Learning Analytics to Scale the Provision of Personalised Feedback.” *British Journal of Educational Technology* 50 (1): 128–38. https://doi.org/10.1111/bjet.12592.

Pasquale, Frank. 2015. *The Black Box Society*. Harvard University Press. http://www.hup.harvard.edu/catalog.php?isbn=9780674368279.

Perrotta, Carlo. 2013. “Assessment, Technology and Democratic Education in the Age of Data.” *Learning, Media and Technology* 38 (1): 116–22. https://doi.org/10.1080/17439884.2013.752384.

Prinsloo, Paul, and Sharon Slade. 2014. “Educational Triage in Open Distance Learning: Walking a Moral Tightrope.” *The International Review of Research in Open and Distance Learning* 15 (4). http://www.irrodl.org/index.php/irrodl/article/view/1881.

Richardson, Rashida, Jason Schultz, and Kate Crawford. 2019. “Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice.” SSRN Scholarly Paper ID 3333423. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3333423.

Roberts, Lynne D., Vanessa Chang, and David Gibson. 2017. “Ethical Considerations in Adopting a University- and System-Wide Approach to Data and Learning Analytics.” In *Big Data and Learning Analytics in Higher Education: Current Theory and Practice*, edited by Ben Kei Daniel, 89–108. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_7.

Rubel, Alan, and Kyle M. L. Jones. 2016. “Student Privacy in Learning Analytics: An Information Ethics Perspective.” *The Information Society* 32 (2): 143–59. https://doi.org/10.1080/01972243.2016.1130502.

Selbst, Andrew D., Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. “Fairness and Abstraction in Sociotechnical Systems.” In *Proceedings of*

the Conference on Fairness, Accountability, and Transparency, 59–68. FAT* '19. New York, NY, USA: ACM. https://doi.org/10.1145/3287560.3287598.
Shum, Simon Buckingham, and Ruth Deakin Crick. 2012. "Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics." In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, 92–101. LAK '12. New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330629.
Siemens, George. 2013. "Learning Analytics: The Emergence of a Discipline." *American Behavioral Scientist* 57 (10): 1380–1400. https://doi.org/10.1177/0002764213498851.
Siemens, George, and Ryan S. J. d. Baker. 2012. "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration." In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, 252–254. LAK '12. New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330661.
Simon, F., M. Kuczera, and B. Pont. 2007. "No More Failures: Ten Steps to Equity in Education. Summary and Policy Recommendations." OECD. http://www.oecd.org/education/school/nomorefailurestenstepstoequityineducation.htm.
The European Parliament and the Council of the European Union. 2016. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance)." *Official Journal of the European Union* OJ (L119): 1–88.
Thompson, Greg, and Ian Cook. 2017. "The Logic of Data-Sense: Thinking through Learning Personalisation." *Discourse: Studies in the Cultural Politics of Education* 38 (5): 740–54. https://doi.org/10.1080/01596306.2016.1148833.
Tsai, Yi-Shan, and Dragan Gašević. 2017. "Learning Analytics in Higher Education — Challenges and Policies: A Review of Eight Learning Analytics Policies." In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 233–242. LAK '17. New York, NY, USA: ACM. https://doi.org/10.1145/3027385.3027400.
Tsai, Yi-Shan, Pedro Manuel Moreno-Marcos, Ioana Jivet, Maren Scheffel, Kairit Tammets, Kaire Kollom, and Dragan Gašević. 2018. "The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics." *Journal of Learning Analytics* 5 (3): 5–20. https://doi.org/10.18608/jla.2018.53.2.
Tsai, Yi-Shan, Pedro Manuel Moreno-Marcos, Kairit Tammets, Kaire Kollom, and Dragan Gaš ević. 2018. "SHEILA Policy Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics." In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 320–329. LAK '18. New York, NY, USA: ACM. https://doi.org/10.1145/3170358.3170367.
Turow, Joseph, Lee McGuigan, and Elena R Maris. 2015. "Making Data Mining a Natural Part of Life: Physical Retailing, Customer Surveillance and the 21st Century Social Imaginary." *European Journal of Cultural Studies* 18 (4–5): 464–78. https://doi.org/10.1177/1367549415577390.
Williamson, Ben. 2015. "Governing Software: Networks, Databases and Algorithmic Power in the Digital Governance of Public Education." *Learning, Media and Technology* 40 (1): 83–105. https://doi.org/10.1080/17439884.2014.924527.
Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30: 75–89.